

Sesgos en estudios epidemiológicos

Mauricio Hernández-Avila, Ph.D.,⁽¹⁾ Francisco Garrido, M.C., M. en C.,⁽²⁾
Eduardo Salazar-Martínez, Dr. en C.⁽¹⁾

Un objetivo en todo estudio epidemiológico, en el que se desea conocer la frecuencia con la que ocurre un evento o estimar la asociación entre un factor de riesgo y una enfermedad, debe ser el de medir y calcular con la mayor precisión y exactitud posibles dichas determinaciones. En otras palabras, la validez del conocimiento derivado de cualquier estudio epidemiológico dependerá, en gran medida, de la ausencia de error y de la capacidad de estimar o predecir el parámetro verdadero en la población blanco. En el contexto de la epidemiología, la validez se refiere a la ausencia de sesgo o error. A lo largo del presente trabajo se hace referencia a dos tipos de validez: a) la validez interna, que se refiere principalmente a los errores cometidos durante el proceso de selección de la población de estudio, durante las mediciones que se realizan en dicha población o a errores ocasionados por la falta de comparabilidad de los grupos estudiados, y b) la validez externa, que se refiere a la capacidad del estudio de generalizar los resultados observados en la población en estudio hacia la población blanco. Es importante notar que la validez externa depende de que exista la interna, es decir, es necesario cumplir con los requisitos necesarios de validez interna para poder extrapolar los resultados, por esta razón en los diferentes estudios epidemiológicos se privilegian acciones que maximizan la validez interna, aun comprometiendo, en cierta medida, la validez externa.

Todo estudio epidemiológico está sujeto a un cierto margen de error, por lo que será muy importante conocer cuáles son sus fuentes principales y los diferentes procedimientos que pueden ser utilizados para minimizar su impacto en los resultados. Los errores se pueden clasificar en dos grandes tipos: los errores no aleatorios o sistemáticos y los errores aleatorios (no sistemáticos); ambos tipos de error, de no controlarse adecuadamente, pueden comprometer la validez del estudio. El error aleatorio (no sistemático) ocurre cuando las mediciones repetidas, ya sean en un mismo sujeto o en diferentes miembros de la población en estudio, varían de manera no predecible, mientras que el error sistemático (no aleatorio) ocurre cuando estas medidas varían de manera predecible y, por lo tanto, se tiende a sobre o subestimar el valor verdadero en medidas repetidas. La analogía que se utiliza para describir ambos conceptos es la práctica de "tiro al blanco" donde el punto medio del objetivo es el valor verdadero en la población blanco y los "disparos" son las diferentes mediciones que se realizan en la población en estudio para estimar dicho valor verdadero. Un buen tirador cuya arma no está bien calibrada apuntará al blanco equivocado, podrá ser muy preciso (todos los disparos dan en el mismo lugar), pero ninguno de ellos da en el blanco correcto. Esto corresponde al error sistemático. Por otra parte, un tirador con mano temblorosa, pero con un arma bien calibrada, estará apuntando al blanco correcto

(1) Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública (INSP), Cuernavaca, Morelos, México.

(2) Centro de Investigación en Sistemas de Salud, INSP, Cuernavaca, Morelos, México.

aun cuando sus disparos no den en el punto medio del blanco seleccionado. Se puede aplicar la analogía descrita en el contexto de un estudio epidemiológico cuyo principal objetivo es determinar las concentraciones de plomo en sangre en la población general. Suponga que en este estudio participan cuatro laboratorios y un laboratorio de referencia en el cual se repite un cierto porcentaje de las determinaciones de plomo. Los resultados se muestran en la figura 1. En el panel A se muestra un laboratorio cuyas determinaciones son válidas y precisas respecto de las mediciones que realiza el laboratorio de referencia (situación ideal) ya que el valor estimado corresponde muy cercanamente al valor real; en el panel B se muestran resultados de un laboratorio cuyas mediciones son válidas pero con precisión intermedia; en el laboratorio del panel C la validez es cuestionable, pero la precisión es aceptable; mientras que en el panel D las mediciones son imprecisas y no válidas. Mientras que en los dos primeros escenarios el parámetro estimado en la población en estudio es el mismo, en ambos laboratorios se estima una media poblacional de 12.2 $\mu\text{g/dl}$, que

corresponde a la media verdadera de 12.2 $\mu\text{g/dl}$, en los paneles C y D se estimaría una media de 15.4 $\mu\text{g/dl}$, y 27.7 $\mu\text{g/dl}$, respectivamente. La utilización de laboratorios con resultados como los que se muestran en los paneles C y D para determinar las concentraciones de plomo en sangre en la población en estudio arrojaría resultados claramente erróneos al sobrestimar el valor real, lo que nos llevaría a concluir equivocadamente que existen concentraciones de plomo en sangre más elevadas que las reales.

En los estudios epidemiológicos analíticos en los que se pone a prueba una hipótesis comparando dos o más grupos de estudio, los errores también pueden ser tanto aleatorios o no sistemáticos como sistemáticos. El error sistemático o sesgo se ha definido como cualquier error diferencial –en relación con los grupos que se comparan– en que se puede incurrir durante el diseño, conducción o análisis del estudio y que invariablemente resulta en una conclusión errónea, ya sea proporcionando una estimación más baja o más alta del valor real de la asociación que existe en la población blanco.

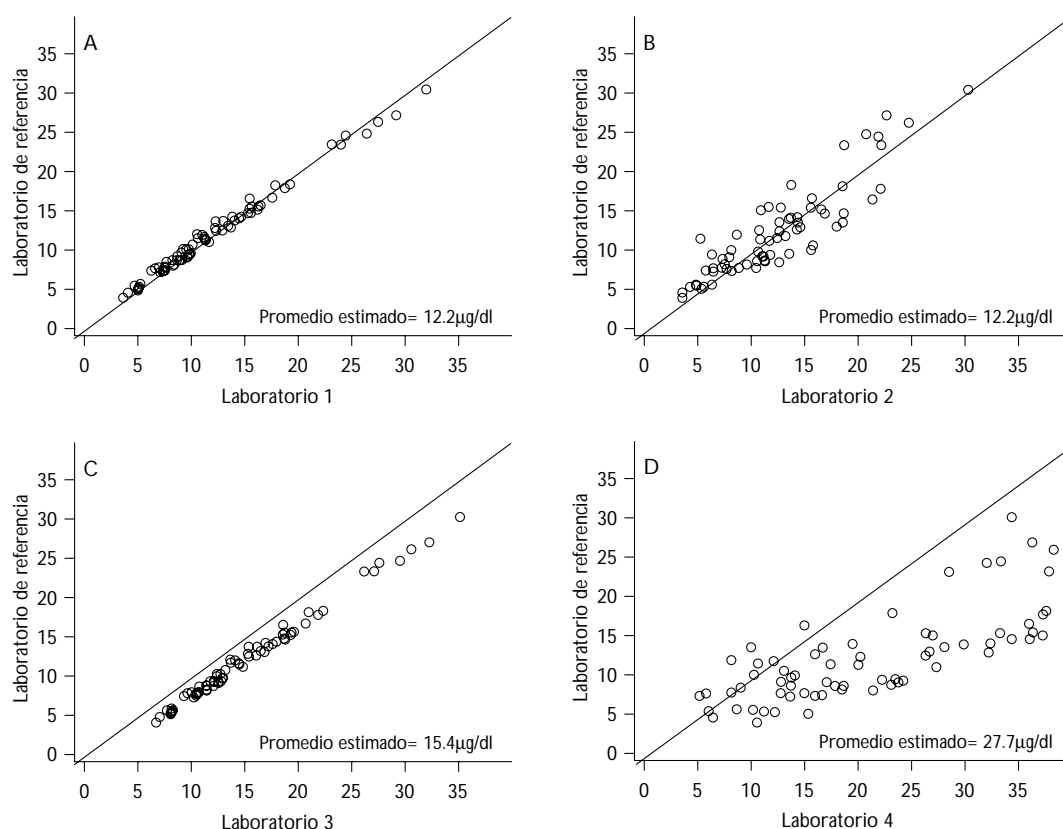


FIGURA 1. VALORES DE PLOMO EN SANGRE REPORTADOS EN CUATRO LABORATORIOS CON DIFERENTES CONDICIONES DE PRECISIÓN Y VALIDEZ

Dependiendo de la etapa del estudio en que se originan, los sesgos que interfieren con la validez interna de un estudio se han clasificado en tres grandes grupos: a) los sesgos de selección, que se refieren a los errores que se introducen durante la selección o el seguimiento de la población en estudio; b) los sesgos de información, que son errores en los que se incurre durante los procesos de medición en la población en estudio, y c) los sesgos de confusión, que se originan por la imposibilidad de asignar la exposición de manera aleatoria en los estudios observacionales y que básicamente se originan por una no comparabilidad de los grupos en estudio. Todo diseño epidemiológico no experimental, en mayor o menor medida, es susceptible de este tipo de sesgos por lo que es un imperativo para los investigadores planear adecuadamente cada una de las etapas de un estudio con el propósito de evitar o disminuir al máximo la posibilidad de incurrir en dichos errores. Sin duda, la etapa más crítica de un estudio corresponde a la del diseño ya que resulta casi imposible corregir *a posteriori* los sesgos introducidos durante esta etapa.

Sesgos de selección

Son errores sistemáticos que se introducen durante la selección o el seguimiento de la población en estudio y que propician una conclusión equivocada sobre la hipótesis en evaluación. Los errores de selección pueden ser originados por el mismo investigador o ser el resultado de relaciones complejas en la población en estudio que pueden no ser evidentes para el investigador y pasar desapercibidas. En este contexto, una posible fuente de sesgo de selección puede ser cualquier factor que influya sobre la posibilidad de los sujetos seleccionados de participar o permanecer en el estudio y que, además, esté relacionado con la exposición o con el evento en estudio.

Los sesgos de selección pueden ocurrir en cualquier estudio epidemiológico, sin embargo, ocurren con mayor frecuencia en estudios retrospectivos y, en particular, en estudios transversales o de encuesta. En los estudios de cohorte prospectivos los sesgos de selección ocurren raramente ya que el reclutamiento y selección de la población en estudio se da antes de que ocurra el evento en estudio, así que se puede suponer que la selección de los participantes se realiza de manera independiente del evento y, en general, la participación en el estudio no puede ser influida por el evento, ya que éste aún no ha ocurrido. En contraste, la permanencia de los participantes en el estudio sí puede ser determinada por el evento, cuando esto

ocurre, y es de diferente magnitud para los grupos expuesto y no expuesto, existirá la posibilidad de que los resultados se vean distorsionados por esta permanencia diferencial. Por esta razón, se recomienda maximizar las tasas de permanencia y seguimiento en los estudios de cohorte.

En los estudios retrospectivos los sesgos de selección pueden ocurrir cuando los participantes potenciales o los investigadores conocen la condición de exposición y/o de enfermedad, y este conocimiento influye diferencialmente la participación en el estudio. Suponga que se llevará a cabo un estudio de una cohorte retrospectiva y que el resultado de este estudio servirá para fijar una posible compensación económica si se logra comprobar que los sujetos que recibieron cierta exposición están en mayor riesgo de desarrollar la enfermedad en estudio. Si esta información fuera del conocimiento de los posibles participantes podría condicionar que los sujetos que se saben expuestos y que desarrollaron la enfermedad en estudio participaran más frecuentemente que aquellos que desarrollaron la enfermedad y que se identifican como no expuestos. La participación diferencial antes mencionada se reflejaría en un déficit de eventos en el grupo de no expuestos, ya que los sujetos no expuestos que desarrollaron la enfermedad no tendrían la misma motivación para participar en el estudio y participarían en menor proporción. Esta participación diferencial condicionaría una subestimación de la frecuencia del evento en el grupo no expuesto y una sobrestimación de la diferencia real.

Los estudios de casos y controles son particularmente susceptibles a este tipo de sesgo ya que en la mayoría de sus aplicaciones se trata de estudios retrospectivos. En el contexto de este tipo de estudios, cualquier factor que influya sobre la probabilidad de selección, ya sea como caso o control, y que a su vez esté relacionado con la exposición en estudio, será una posible fuente de sesgo de selección. Este tipo de sesgo se ha propuesto como una de las explicaciones a la asociación reportada entre el consumo de café y el cáncer de páncreas por McMahon y colaboradores.¹ En este estudio los casos de cáncer de páncreas se seleccionaron de diferentes hospitales, mientras que los controles fueron seleccionados del mismo grupo médico de donde se había originado el diagnóstico de los casos. Al elegir este mecanismo de selección de controles los investigadores incluyeron como controles a diferentes sujetos con patología gastrointestinal, dado que la mayoría de los casos de cáncer de páncreas se habían notificado por el servicio de gastroenterología. Así, los controles representan un sector

de la población con un consumo menos frecuente de café, lo que pudo haber condicionado una subestimación de la prevalencia de exposición en el grupo control y, por lo tanto, una asociación espuria entre el consumo de café y el cáncer de páncreas.

La detección diferencial es un tipo particular de sesgo de selección, se origina cuando la prueba diagnóstica para detectar el evento se realiza con mayor frecuencia en el grupo expuesto. Un ejemplo de este sesgo se presentó en un estudio de casos y controles en el que se observó una fuerte asociación entre el uso de estrógenos de remplazo y el cáncer de endometrio.² Los casos en este estudio no representaban adecuadamente los casos originados en la población base, ya que las mujeres con síntomas sugestivos de cáncer de endometrio, y que eran usuarias de estrógenos, eran hospitalizadas para diagnóstico con mayor frecuencia que las mujeres con los mismos síntomas pero que no eran usuarias de estrógenos. Esta conducta de los médicos generó una sobrerselección de casos expuestos a estrógenos, lo que ocasionó una asociación falsa positiva entre el uso de estrógenos y el riesgo de cáncer de endometrio en el estudio.

La no respuesta, por parte de los participantes en un estudio, puede también introducir sesgo de selección, siempre y cuando esté relacionada con la exposición o el evento en estudio, es decir, que la tasa de participación sea diferente para expuestos y no expuestos en los estudios de cohorte, o para casos y controles en el contexto de un estudio de casos y controles. Sin embargo, no siempre la falta de respuesta se asocia con un sesgo de este tipo. Por ejemplo, si en un estudio de casos y controles 10% de los casos elegidos y 30% de los controles rehusan participar no se incurrirá en sesgo de selección, a menos que la proporción de expuestos en los casos y en los controles que no aceptaron participar sea diferente a la del grupo respectivo. La no respuesta es un problema que tiene mayor relevancia para la validez externa.

Los estudios transversales, y de casos y controles que se basan en casos existentes (prevalentes), presentan importantes limitaciones relacionadas con los sesgos de selección, en particular cuando la enfermedad en estudio tiene una alta letalidad cercana al diagnóstico inicial ya que los casos existentes tienden a sobrerrepresentar a los sujetos con cursos más benignos de la enfermedad. Si el factor en estudio se asocia con la letalidad, la medida de efecto derivada de un estudio de prevalencia será sesgada, dado que los casos en estudio corresponden a sobrevivientes de la enfermedad, por lo que representarán desproporcionadamente a los casos no expuestos.

Sesgos de información

El sesgo de información se refiere a los errores que se introducen durante la medición de la exposición, de los eventos u otras covariables en la población en estudio, que se presentan de manera diferencial entre los grupos que se comparan, y que ocasionan una conclusión errónea respecto de la hipótesis que se investiga. Una posible fuente de sesgo de medición puede ser cualquier factor que influya de manera diferencial sobre la calidad de las mediciones que se realizan en los grupos expuesto y no expuesto en el contexto de los estudios de cohorte o entre los casos y controles en el contexto de los estudios de casos y controles.

Es importante mencionar que aunque prácticamente no existen procedimientos libres de error de medición, no todos los errores de medición son fuente de sesgo de información. Es conveniente recordar que los errores de medición pueden ser no diferenciales (aleatorios) cuando el grado de error del instrumento o técnica empleada es el mismo para los grupos que se comparan y diferenciales (no aleatorios) cuando el grado de error es diferente para los grupos estudiados, el sesgo de información se refiere particularmente a este último tipo. Para comprender mejor la diferencia entre ambos errores de medición se analiza el siguiente ejemplo. En un estudio hipotético de casos y controles para evaluar la asociación entre tabaquismo e infarto agudo del miocardio, los casos se identificaron en las salas de urgencias, del Instituto Mexicano del Seguro Social al momento de su ingreso y los controles fueron seleccionados al azar entre los derechohabientes que son vecinos de cada caso. Suponga que, en un primer escenario, la exposición al tabaco se evalúa determinando la presencia de un marcador biológico de exposición que se mide en sangre. En este primer escenario, puede existir cierto grado de error en las determinaciones del biomarcador en sangre, sin embargo es posible suponer que el error es similar para los dos grupos, por lo que se considera no diferencial o aleatorio. En contraste, si la exposición al tabaco se hubiera evaluado mediante un cuestionario, la calidad de la información dependería, en parte, de la memoria de los participantes. Si los casos, dado que sufrieron el evento, tuvieran un estímulo mayor para recordar o participar entonces la calidad de la información sería mejor en este grupo que la que se podría obtener en el grupo control, situación que introduciría un error diferencial o no aleatorio. En general, el impacto de este último tipo de error es difícil de predecir ya que puede subestimar o sobrestimar la asociación real, a diferencia del error

aleatorio que en general tiende a subestimar las asociaciones reales.

En el cuadro I se ofrece una clasificación general de las diferentes fuentes de error de medición. Es importante hacer notar que durante el diseño de un estudio se deben establecer todas aquellas características que se habrán de medir para responder una pregunta de investigación. Una de las dificultades con las que se topa el investigador en esta etapa se deriva de la brecha o diferencia que se establece entre una entidad conceptual de una variable y su definición empírica que es, finalmente, la forma en que será medida. Esta diferencia estriba en la imposibilidad, la mayor parte de las ocasiones, de medir directamente la característica de interés. Por ejemplo, la exposición involuntaria al humo del tabaco podría ser definida teóricamente como "la exposición de las células del tejido pulmonar a agentes carcinogénicos que resulta de la inhalación del humo del tabaco proveniente de las emisiones de fumadores activos", en la práctica esta definición teórica se tendría que medir mediante la utilización de algún indicador cercano de dicha condición, tal como "tiempo que pasa el sujeto de estudio junto a personas fumadoras". Estas definiciones empíricas son utilizadas y plasmadas por los epidemiólogos y por investigadores de otras disciplinas, en los instrumentos de medición que se elaboran *ex profeso* para un estudio y que pueden ser medidas ya sea por un cuestionario o un biomarcador. En este último ejemplo, la exposición involuntaria se podría también medir mediante la determinación de cotinina (un metabolito de la nicotina) en la sangre de los sujetos en estudio. Mientras que las estimaciones de exposición realizadas mediante un cuestionario estarán sujetas a diferentes tipos de error como lo serían la memoria de los participantes y la comprensión de las

preguntas, en el caso del biomarcador, serán fuentes de error los procedimientos utilizados para obtener la muestra, el transporte y conservación de la misma, así como los diferentes procedimientos de laboratorio que se necesiten utilizar para su medición.

Para ilustrar el impacto que pueden tener los diferentes tipos de error de información en la estimación de las medidas de efecto, es necesario mencionar brevemente los conceptos de sensibilidad y especificidad de los instrumentos de medición que se utilizan. La sensibilidad de un cuestionario o prueba diagnóstica mide la habilidad del instrumento para clasificar correctamente a los sujetos que tienen verdaderamente la condición en estudio. Así, un cuestionario para identificar fumadores involuntarios que tiene una sensibilidad de 90%, podrá identificar correctamente a 90% de los sujetos que están expuestos al humo del tabaco, mientras que 10% serán clasificados erróneamente como no fumadores involuntarios. En contraste, la especificidad indica la habilidad del instrumento para clasificar correctamente a los sujetos que no tienen la condición de estudio. Para este ejemplo, una especificidad de 99% indicaría que se clasificaría adecuadamente a 99% de los sujetos y que restaría un 1% que sería clasificado como expuesto involuntariamente al tabaco, cuando en realidad no lo está. La sensibilidad y especificidad de nuestros instrumentos puede ser evaluada en estudios de validación cuando se cuenta con un estándar de oro, el cuadro II indica cómo se podrían estimar estos parámetros.

Los sujetos de estudio se pueden clasificar erróneamente con relación a la exposición o al evento, como se mencionó, el error puede ser de dos tipos: no diferencial y diferencial.

Los estudios epidemiológicos incurren en un error de clasificación no diferencial de la exposición cuando ésta se realiza por igual en los grupos que se

Cuadro I
CLASIFICACIÓN DEL ERROR DE MEDICIÓN SEGÚN SU ORIGEN

- a) El observador
- b) Sistema de medición
- c) Los sujetos de estudio
 - Memoria
 - Entrenamiento
 - Fatiga
- d) El instrumento
- e) Errores en las variables proxy
- f) El procesamiento de datos
 - Errores de codificación
 - Formulación errónea de modelos estadísticos
- g) Errores que dependen del tiempo

Cuadro II
ESTIMACIÓN DE SENSIBILIDAD Y ESPECIFICIDAD PARA UNA VARIABLE DICOTÓMICA

Clasificación de la exposición o de la enfermedad por el instrumento de medición	Clasificación verdadera de la exposición o de la enfermedad	
	Sí	No
Sí	a	b
No	c	d

Sensibilidad = $\frac{a}{a+c}$; Especificidad = $\frac{d}{b+d}$

comparan. Es decir, la condición de enfermedad de los sujetos de estudio no influye en la clasificación de los mismos respecto de la exposición. En esta situación, siempre y cuando el error no sea de gran magnitud, la medida de efecto estimada con error aleatorio tiende a subestimar la asociación real en la población blanco, por lo que se podría suponer que el valor estimado en la población en estudio, bajo el supuesto de que el error es aleatorio, tiende o se aproxima al valor de no asociación. La dificultad para medir correctamente la exposición involuntaria al humo del tabaco puede servir para ilustrar este tipo de error. Dado que no es factible medir directamente la cantidad de humo que inhala un no fumador se podría utilizar un monitor personal que registre en forma acumulada la concentración de nicotina ambiental a la que el sujeto estuvo expuesto durante el tiempo relevante para el evento en estudio. Supongamos un estudio longitudinal para examinar la asociación entre enfermedad respiratoria y exposición involuntaria al humo del tabaco. En este caso, el error de clasificación –no diferencial– en la medición de la exposición estaría dado por la habilidad del sujeto de llevar consigo el monitor, del tipo de ropa que utilice y de la conservación y trato que dé al instrumento de medición. En este caso se podría suponer que el grado de error será el mismo para los sujetos con y sin exposición al humo de tabaco y que el impacto de dicho error será el de subestimar cualquier asociación real que exista en la población blanco.

Otro ejemplo que permite ilustrar el contexto en el que ocurre el error de medición aleatorio o no diferencial es cuando la exposición en estudio no considera el periodo de riesgo relevante en la ocurrencia de una enfermedad. Si se estudia una enfermedad que tiene un periodo de latencia prolongado, como es el caso de la gran mayoría de las enfermedades crónicas, y el factor de exposición que se cree asociado con dicha enfermedad se mide en una sola ocasión, sin tomar en consideración la variabilidad de dicho factor durante el tiempo, se incurrirá en un error de clasificación no diferencial. De manera más explícita, se puede decir que si se desea medir la asociación entre el consumo de colesterol y enfermedad coronaria para lo cual medimos la ingesta de colesterol en un solo día, no se puede asegurar que se haya medido correctamente la exposición dado que la ingesta varía considerablemente a través del tiempo en un mismo sujeto y, por lo tanto, una sola medición no reflejaría adecuadamente la ingesta de colesterol durante el periodo de riesgo relevante.

A continuación se ofrece un ejemplo hipotético para ilustrar el efecto del error aleatorio o no diferen-

cial. En el cuadro III se resumen los datos de un estudio de seguimiento por 12 años en hombres de 45-64 años ($n=7\,221$) en el que se relacionó el ejercicio físico con el riesgo de enfermedad coronaria. Los expuestos fueron quienes no hacían ejercicio físico o lo hacían moderadamente. Asumamos que la medida de efecto en este primer escenario se estima sin error aleatorio y corresponde a una razón de incidencia acumulada de 1.41.

Supongamos ahora que la determinación del evento se realiza con una prueba diagnóstica que tiene una sensibilidad de 0.80 y especificidad de 1.0 y que la prueba se aplica de igual manera en expuestos y en no expuestos. Los resultados se presentan en el cuadro III-A, donde se puede observar que existe un subestimación del riesgo real, la razón de incidencia acumulada disminuye de 1.41 a 1.38. Sin embargo, si se supone una especificidad de 98%, entonces el error es de mayor impacto ya que la razón de incidencia estimada bajo este supuesto sería de 1.28.

Cuadro III
RELACIÓN ENTRE EJERCICIO FÍSICO Y RIESGO
DE ENFERMEDAD CORONARIA,
ESTIMADA SIN ERROR ALEATORIO

Evento coronario	Realiza ejercicio físico		Total
	No	Sí	
Sí	299	107	406
No	4 503	2 312	6 815
Total	4 802	2 419	7 221

Razón de incidencia acumulada = $299/4\,802 + 107/2\,419 = 1.41$

Cuadro III-A
RELACIÓN ENTRE EJERCICIO FÍSICO Y RIESGO
DE ENFERMEDAD CORONARIA, ESTIMADA
CON UNA PRUEBA DIAGNÓSTICA CUYA SENSIBILIDAD
ES DE 0.80 Y ESPECIFICIDAD DE 1.0

Enfermedad	Realiza ejercicio físico		Total
	No	Sí	
Sí	239	87	406
No	4 563	2 332	6 815
Total	4 802	2 419	7 221

Razón de incidencia acumulada = $239/4\,802 + 87/2\,419 = 1.38$

239 se deriva de 299 casos de los cuales únicamente 80% (299×0.80) fueron clasificados correctamente como eventos
4 563 corresponde a 4 503 sujetos sin eventos y a 60 sujetos con evento que fueron clasificados erróneamente como sanos

Como se mencionó, cuando el error de medición se realiza de manera diferencial entre los grupos que se comparan, entonces se introduce un sesgo de información que, a diferencia del error aleatorio que condiciona una subestimación, puede introducir error en cualquier dirección, condicionando ya sea una sobre o una subestimación de la asociación real. Existen varias situaciones que se pueden asociar con este sesgo. En la etapa de diseño del estudio se puede favorecer la introducción del sesgo cuando se decide recoger la información sobre la exposición después que se han identificado los casos de la enfermedad de interés y la presencia de la enfermedad modifica el reporte de la exposición (exagerándola o minimizándola). En el caso del estudio de exposición involuntaria al humo del tabaco antes referido, si los investigadores utilizaran un cuestionario como instrumento para la evaluación de tipo retrospectivo, no es difícil suponer que las madres cuyos hijos tuvieron problemas respiratorios tratarán de recordar con mayor detalle y exactitud la exposición al humo del tabaco comparadas con aquellas madres cuyos hijos permanecieron sanos durante el periodo de estudio; o que las madres fumadoras cuyos hijos tuvieron enfermedad respiratoria informarían una menor exposición como un mecanismo compensatorio de culpa. En ambas situaciones es claro que se puede incurrir en sesgos de información diferencial.

El sesgo de clasificación diferencial de la exposición ocurre también con frecuencia durante la fase de campo. El sesgo del entrevistador, un caso de sesgo de información, se puede introducir cuando el personal de campo conoce la condición de enfermo o de no enfermo del entrevistado y si, además, tiene conocimiento de las hipótesis u objetivos del estudio. Ante este hecho, el entrevistador puede hacer más énfasis en aquellas preguntas que miden la exposición o bien inducir las respuestas de los entrevistados mediante lenguaje corporal o la aportación de información adicional sobre dichas preguntas. En este sentido es recomendable realizar un buen entrenamiento de los entrevistadores y mantenerlos ciegos tanto en lo que respecta a la hipótesis del estudio como a la condición de caso-control o de expuesto-no expuesto de los participantes del estudio.

El uso de informantes sustitutos (proxy) puede también dar lugar a un sesgo de clasificación diferencial de la exposición. Como se mencionó, con cierta frecuencia los estudios epidemiológicos deben recurrir a informantes que no son los sujetos de estudio ya sea por alguna incapacidad de éstos o bien porque ya fallecieron. Dependiendo del tipo de relación que tiene o tuvo el informante sustituto con el

sujeto de estudio, las respuestas que éstos proporcionan sobre la exposición de interés pueden oscilar entre la negación o la exageración. Es probable que si la exposición se refiere a una característica considerada negativa y el informante sustituto es la esposa(o) del sujeto de estudio, las respuestas que se proporcionen se pueden modificar con base en la relación existente entre ambos.

Para ilustrar el impacto del error diferencial, supongamos ahora que la relación descrita en el ejemplo anterior entre ejercicio y enfermedad coronaria se evalúa mediante un estudio de casos y controles. En este estudio se identificarían 406 casos y para cada caso se seleccionarían dos controles, los resultados hipotéticos se presentan en el cuadro IV. La razón de momios (RM) estimada del estudio es de 1.40.

Ahora, supongamos que los casos, dado que sufrieron el evento coronario, hacen un mejor esfuerzo por recordar y reportar el patrón de ejercicio que realizaban y que la sensibilidad con que informan la exposición es de 95%, mientras que los controles reportan con una sensibilidad de 80%. Los resultados se presentan en el cuadro IV-A. En este escenario, el sesgo introducido por la menor calidad de reporte en los

Cuadro IV
RELACIÓN ENTRE EJERCICIO FÍSICO Y RIESGO
DE ENFERMEDAD CORONARIA, ESTIMADA
SIN ERROR DIFERENCIAL

Exposición	Hace ejercicio		Total
	No	Sí	
Casos	299	107	406
Controles	540	272	812
Total	839	379	1 218

Razón de momios= $299 \times 272 / 540 \times 107 = 1.40$

Cuadro IV-A
RELACIÓN ENTRE EJERCICIO FÍSICO Y RIESGO
DE ENFERMEDAD CORONARIA, ESTIMADA A PARTIR DE
UN INFORME DE LA EXPOSICIÓN CON SENSIBILIDAD
DE 95% EN LOS CASOS Y DE 80% EN LOS CONTROLES

	Ejercicios	
	No	Sí
Casos	304	102
Controles	594	218

Sensibilidad en casos= 0.95 en controles= 0.80
Especificidad en casos= 1.0 en controles= 1.0

Razón de momios= $304 \times 218 / 594 \times 102 = 1.09$

casos condiciona una subestimación de la razón de momios.

Supongamos ahora que los casos tienden a subestimar el reporte de ejercicio y que la sensibilidad en este grupo es de 70% mientras que en el grupo control el cuestionario utilizado clasifica adecuadamente la exposición en 90% de los sujetos. Bajo los supuestos en este escenario, el sesgo diferencial de reporte condicionaría una sobrestimación del efecto de no hacer ejercicio (cuadro IV-B). Como se puede apreciar, el efecto puede condicionar desviaciones hacia cualquier dirección, por esta razón es necesario evitar en la medida de lo posible introducir errores diferenciales.

Sesgos de confusión

Todos los resultados derivados de estudios observacionales están potencialmente influenciados por este tipo de sesgo. El sesgo de confusión puede resultar en una sobre o subestimación de la asociación real. Existe sesgo de confusión cuando observamos una asociación no causal entre la exposición y el evento en estudio o cuando no observamos una asociación real entre la exposición y el evento en estudio por la acción de una tercera variable que no es controlada. Esta(s) variable(s) se denomina(n) factor(es) de confusión o confusor(es). Los resultados de un estudio estarán confundidos cuando los resultados obtenidos en la población en estudio apoyan una conclusión falsa o espuria sobre la hipótesis en evaluación, debido a la influencia de otras variables, que no fueron controladas adecuadamente ya sea durante la fase de diseño o de análisis. En este contexto, son fuente posible de sesgo de confusión cualquier variable asociada con la exposición que, además, esté causalmente asociada con el

evento en estudio y que se encuentre distribuida de manera diferencial entre los grupos que se comparan, ya sea entre expuestos y no expuestos en el contexto de los estudios de cohorte o entre casos y controles en el ámbito de los estudios de casos y controles.

En los estudios observacionales el sesgo de confusión se puede entender como un problema de comparabilidad cuyo origen está ligado a la imposibilidad de realizar una asignación aleatoria de la exposición en los sujetos de estudio. El objetivo de la asignación al azar de los tratamientos (de la exposición) en los estudios experimentales es lograr la formación de grupos homogéneos en lo que se refiere a todas las características que puedan influir en el riesgo de desarrollar el evento (edad, sexo, masa corporal u otras características que no se puedan medir), lo que se busca lograr es que los grupos sean similares en todo excepto en la exposición que se busca evaluar.

Analicemos el siguiente ejemplo, cuando Doll y Hill publicaron sus primeros resultados derivados del seguimiento de cerca de 40 000 médicos y reportaron que el cáncer de pulmón era considerablemente más frecuente en los fumadores,^{3,4} las críticas no se hicieron esperar. Se argumentó que la asociación observada por estos investigadores entre el tabaquismo y el cáncer de pulmón se debía muy probablemente a la acción de una tercera variable, a una susceptibilidad genética que predisponía tanto al cáncer de pulmón como al gusto por el tabaco, con este argumento se descartaba el posible efecto carcinogénico del humo de tabaco. La crítica anterior implica que el gusto por fumar no se da al azar sino que coexiste un factor genético y que este factor es la verdadera causa de cáncer de pulmón. Para poder separar el efecto del cigarrillo del de la susceptibilidad genética tendríamos que realizar un experimento, se podrían seleccionar 40 000 individuos y asignarlos aleatoriamente a dos tratamientos experimentales: fumadores (expuesto) y no fumadores (no expuesto). Al asignarlos al azar se puede suponer que la distribución del factor genético que causa cáncer de pulmón en los grupos formados mediante la aleatorización será la misma, es decir, si la aleatorización se realizó de manera correcta –en promedio– el porcentaje de sujetos genéticamente susceptibles al cáncer de pulmón sería la misma en ambos grupos, por lo tanto, si observáramos diferencias en la ocurrencia de cáncer de pulmón entre los fumadores y no fumadores ésta no podría ser atribuida al factor genético ya que éste afectaría de igual manera a los dos grupos. Sin embargo, es claro que este tipo de experimento sería imposible de realizar.

A diferencia de los estudios experimentales, en que los sujetos en estudio son asignados al azar a los

Cuadro IV-B
RELACIÓN ENTRE EJERCICIO FÍSICO Y RIESGO DE ENFERMEDAD CORONARIA, ESTIMADA A PARTIR DE UN INFORME DE LA EXPOSICIÓN CON SENSIBILIDAD DE 70% EN LOS CASOS Y DE 90% EN LOS CONTROLES

	Ejercicios	
	No	Si
Casos	331	75
Controles	563	249

Sensibilidad en casos= 0.70 en controles= 0.90

Especificidad en casos= 1.0 en controles= 1.0

Razón de momios= 1.95

grupos experimentales y, por lo tanto, los posibles confusores quedan igualmente distribuidos entre los grupos en contraste, en los estudios observacionales los sujetos reciben la exposición por muy diferentes motivos que dependen importantemente de patrones culturales y socioeconómicos. Entonces ¿qué es una variable confusora o confusor? Se trata de variables que están asociadas con el evento en estudio de manera causal, es decir, son factores de riesgo para el evento en estudio, además esta asociación debe ser por un mecanismo causal diferente al que actúa la exposición en estudio y debe ser independiente de esta misma. Esto último quiere decir que el confusor debe ser un factor de riesgo tanto para el grupo expuesto como para el no expuesto. Los sesgos de confusión en estudios de cohorte ocurren principalmente cuando los grupos expuesto y no expuesto no son comparables. Por ejemplo, si fuera de interés evaluar en un estudio de cohorte el efecto del consumo de café sobre el riesgo de infarto del miocardio y se establecieran dos grupos de contraste, uno conformado por bebedores de café y otro por no bebedores de café, se tendría que considerar el tabaquismo como un posible confusor ya que está asociado con el consumo de café y es un factor de riesgo para infarto de miocardio. En este estudio hipotético ignorar las diferencias en el consumo de tabaco podría hacernos concluir que existe una falsa asociación positiva entre el consumo de café y el riesgo de infarto del miocardio. En el contexto de los estudios de casos y controles ocurre cuando hay diferencias entre casos y controles en relación con una tercera variable, lo que puede ocurrir simplemente por la existencia de confusión en la población fuente o ser generada por el mismo proceso de selección.

¿Qué podemos hacer para prevenir la ocurrencia de sesgos de confusión? En el cuadro V se listan los diferentes procedimientos tanto en la fase de diseño como en la de análisis que nos permiten prevenir o corregir el efecto del sesgo de confusión, como se puede deducir de los diferentes procedimientos utilizados, el

Cuadro V
ESTRATEGIAS PARA EL CONTROL DE LA CONFUSIÓN EN
LOS ESTUDIOS EPIDEMIOLÓGICOS

Fase	Estrategia	Efecto
Diseño	Aleatorización	Permite que las variables se distribuyan similarmente en los grupos de estudio haciéndolos comparables en todo, excepto en la variable de exposición
	Restricción	Limita la participación en el estudio a sujetos que son similares respecto a la variable de confusión
	Pareamiento	Iguala en el proceso de selección a los grupos de comparación con relación a los factores de confusión
Análisis	Estandarización	Permite comparar los grupos de estudio si la distribución del confusor fuera la misma en ambos
	Estratificación	Estima la medida de efecto en subgrupos que son similares con relación a los factores de confusión
	Modelos multivariados	Estima el efecto de la exposición, manteniendo constantes los valores del factor confusor

común denominador de todos es hacer los grupos de contraste lo más comparable posible con relación a variables externas, el objetivo es lograr que la única diferencia entre los grupos sea la característica en estudio.

Referencias

1. MacMahon B, Yen S, Trichopoulos D *et al*. Coffee and cancer of the pancreas. *N Engl J Med* 1981; 304:630-633.
2. Feinstein HR, Horowitz RI. A critique of the statistical evidence associating estrogens with endometrial cancer. *Cancer Res* 1978;38: 4001-4005.
3. Doll R, Hill AB. Mortality in relation to smoking: Ten years' observations of British doctors. *BMJ* 1964;1:1399-1410.
4. Doll R, Hill AB. Mortality in Relation to Smoking: Ten years' observations of British doctors (Concluded). *BMJ* 1964;1:1460-1467.